Division of Labour in Collaborative Information Seeking: Current Approaches and Future Directions

Ryan Kelly
University of Bath
Bath, BA2 7AY, United Kingdom
r.m.kelly@bath.ac.uk

Stephen J. Payne
University of Bath
Bath, BA2 7AY, United Kingdom
s.j.payne@bath.ac.uk

ABSTRACT
It is now recognised that people often collaborate when completing information seeking tasks, and a number of specialised tools and systems have been developed to support such behaviour. Such systems often allow for distribution of search results among collaborators. The goal of this division of labour is to enable concurrent work while also preventing redundancy in results distribution. This paper consolidates prior work on division of labour in collaborative information seeking systems by reviewing four approaches to creating division of labour. We then briefly describe our own research before laying out some future directions, which we hope will promote discussion at the workshop.

Keywords
Collaborative information seeking; Collaborative information retrieval; Division of labour; Redundancy

Categories and Subject Descriptors
H.5.m. [**Information Interfaces and Presentation (e.g. HCI)**]: Miscellaneous

General Terms
Human Factors; Economics; Experimentation

1. INTRODUCTION
Recent research has shown that information seeking often occurs among groups of searchers with a shared information need. Such collaborative information seeking (CIS) has been observed across a variety of domains [3, 15, 19] and research has suggested that CIS might be supported through specially designed computational tools, e.g. [7, 16, 22]. However, in order for CIS to be synergistic [21], researchers have argued that CIS technologies must support the processes that underpin collaboration more generally [6, 20, 21]. One such process is *division of labour* (DoL), which refers to the act of breaking up a task such that work is distributed across members of a group [6]. Divided labour has long been a topic of concern for computer-supported cooperative work (e.g. [4, 14]) and a series of authors have argued that a key design goal for CIS systems lies in supporting DoL among searchers [5, 6, 16, 18, 22].

In this paper, we aim to consolidate prior work on DoL in order to set an agenda for future studies of division of labour in CIS systems and scenarios. We begin by using an expanded version of Golovchinsky & Pickens' taxonomy of mediation in CIS [9] to delineate existing CIS tools according to their in-built support for division of labour. We review work on role-based [23], algorithmic [5], communicative [22] and user interface [16] approaches. We assess the relative merits of each approach, and our review allows us to unpack a variety of issues associated with each. We then discuss areas that demand exploration in future work. For the purposes of this paper, our focus is on situations of Web-based CIS where all parties explicitly work together, either synchronously or asynchronously, to intentionally satisfy a shared information need [8]. Such situations are typified by mutual awareness of involvement, and we therefore exclude filtering or recommendation tools that utilise prior searches from anonymous 'collaborators'.

2. DIVISION OF LABOUR IN CIS
Division of labour has been recognised as an important aspect of CIS for a number of reasons. First, an effective division of labour policy facilitates concurrent work among searchers engaged in synchronous information seeking activities [6, 15]. Second, effective DoL allows searchers to prevent redundancy. In the context of CIS, redundancy can prove particularly problematic in that searchers working in parallel may effect similar queries, leading to the retrieval of similar results [5]. This may be of further issue when searching for information that is constrained or indexed by a small number of resources, and thus failure to properly coordinate activity could lead to considerable duplication of effort when conducting collaborative search. In response to these concerns, it has generally been argued that CIS tools should support division of labour by including methods of distributing work that allow searchers to increase their coverage of a topic area while also minimising redundancy [5, 6, 16, 17, 18].

One way of preventing redundancy in CIS is to explicitly mediate work so as to avoid repetition, redundancy, and breakdowns. Although a variety of approaches to resolving this problem have been proposed within the literature, there has been relatively little isolated discussion of the relative strengths of each, and no research has attempting to compare and contrast the various approaches. Our first aim here is to consolidate these issues such that future research will be better placed to target salient problems. To provide a parsimonious discussion, we classify prior approaches by adapting the taxonomy of mediation styles offered

Appeared in CSCW '13 Workshop on Collaborative Information Seeking: Consolidating the Past, Creating the Future.

by Golovchinsky & Pickens [9]. They describe three ways in which organisation of CIS can be achieved: *communicative* mediation, *user interface* mediation, and *algorithmic* mediation. Based on more recent work [11, 23], we further augment Golovchinsky & Pickens' taxonomy with *role-based* division of labour in CIS. In the following sections we classify prior work beneath this taxonomy by focusing on the *in-built* support for division of labour in CIS tools.

2.1 Communicative Approaches

The simplest form of coordinating divided labour is through communication, where people exchange information about their search activity in order to avoid redundancy and coordinate work. Prior work has documented such discussion during CIS [20] and has also identified methods employed by searchers to organise information seeking. For example, Morris [15] identified the *divide-and-conquer* strategy, where coworkers explicitly coordinated their activities by communicating about the use of different search engines or keywords. Morris et al. [17] also observed how coworkers divided labour by verbally negotiating individual rights and responsibilities during the use of a tabletop search interface.

There are several CIS systems that rely solely on communicative mediation for DoL. *Coagmento* [22] invites users to organise their activities as they see fit through the use of synchronous chat and notifications. *SearchTeam*[1] provides several ways of filtering results, e.g. by images, videos or books, but searchers must communicate to explicitly partition the information space to prevent redundancy. *Results Space* [2], a system intended to support asynchronous CIS, provides no in-built DoL capabilities – instead, the system includes a number of awareness mechanisms that allow searchers to see earlier search trails. The rationale here is that users will be able to make judgements about where to look next based on the histories of their collaborators.

One benefit of communicative mediation is that it offers high user freedom. Searchers can organise work as they choose and are not constrained by any formal structures imposed by the system or its designer. The downside, however, is that this approach places the greatest burden on collaborators in terms of articulation work; that is, the 'work required to organise work'. If this cost is too high, searchers may fail to communicate and run the risk of repeated redundancy. Gonzalez-Ibañez et al. [10] also note that communicative mediation is limited by what the individuals know, do, and agree upon. In other words, searchers working on an unfamiliar topic can only communicate about things they already know, meaning that their search behaviour and division of labour policies exist as a function of the group's existing knowledge.

2.2 User Interface Approaches

A number of systems include support for division of labour controlled at the user interface (UI). *SearchTogether* [16], for example, allows collaborators to distribute results in various ways. "Split Search" sends an individual user's query to a search engine and divides the results among all online group members in a round-robin fashion. In contrast, "Multi-Engine Search" assigns results from a single search engine to each collaborator, i.e., one participant receives all results from Bing, another from Google, and so on. *WeSearch* [17] allows searchers to divide individual web pages into 'clips' which facilitate division of labour through concurrent review of separate sections [17].

[1]http://www.searchteam.com

One positive aspect of mediating division of labour at the user interface is that use of such features is not compulsory. Structured support remains available, yet users can also organise their own DoL policies if they choose. However, this gets at the primary problem with UI mediation: many studies have shown that users simply don't use the features provided. For example, separate evaluations of SearchTogether and WeSearch found that UI-level support was neglected in favour of DoL policies negotiated by the group [16, 17]. On the one hand, this suggests that researchers still need to find out what does and does not work with respect to UI mediation. On the other, we note that most evaluations of CIS tools tend to be short-term affairs that utilise artificial tasks to evaluate and identify problems with a particular design (as in [17]). It is certainly possible that users would become accustomed to the use of UI-level tools for DoL if given more time, in turn suggesting a need for more longitudinal and field studies of CIS tools.

One other issue raised by UI-level tools is that of *control*; in particular, who has control over, rights and responsibilities for DoL features and policies. In their evaluation of SearchTogether, Morris & Horvitz [16] found that many participants chose not to use automatic division of labour because only one member of the team had control over the process. This power imbalance ran contrary to the group's collaborative ethos, demonstrating that even seemingly innocuous design decisions can have profound effects on the way a particular system is perceived by its users.

2.3 Algorithmic Approaches

Algorithmic mediation involves the use of data about search activity that is collected by the system and re-used to enhance information seeking [18]. *Cerchiamo* [7] contains a specialised algorithm that "mediates document retrieval at the system layer by allowing independently issued queries to be merged into a core results set, allowing users to collaborate on a corpus of documents without disturbing their partner's work" [18]. The intention is to allow people to work at their own pace but still be influenced in real-time by their partners' search activities. More recently, Foley & Smeaton [5] used algorithmic mediation to enhance division of labour by suppressing the retrieval of documents that had already been evaluated by at least one member of the CIS team. The idea is that "redundant documents in the ranked list are replaced with new unseen documents, [enabling] the group to cover a greater amount of the collection over the course of the search" ([5], p. 768).

One benefit of algorithmic mediation is that coordination costs are very low – by delegating division of labour to the CIS system, collaborators are free to get on with the business of actually finding information. Yet there are also at least two problems with this approach. First, the removal of documents assumes perfect performance in terms of information triage, which refers to the identification of relevant and non-relevant sources by information seekers [1]. It is possible that searchers with low domain knowledge may mistakenly mark relevant material as non-relevant. Second, removing documents may impede the sensemaking process of future collaborators, in the sense that knowing what is *not* relevant may be as important as knowing what is. One other issue is that work on algorithmic approaches has focused on precision-oriented tasks, yet in a task like travel planning, where answers are satisfactory rather than 'correct', users may wish to continually revisit search results in order to reach a final decision with their collaborators. In such circumstances, the removal of certain results could prove detrimental.

2.4 Role-Based Approaches

Finally, a number of CIS studies have explored role assignment as a means of enforcing DoL. The aforementioned system *Cerchiamo* assigns users to the complementary roles of prospector and miner. The prospector spends time finding sources, while the miner evaluates the findings and provides feedback on their relevance [7]. In this way, the role of a particular collaborator determines his or her involvement in the project. Shah et al. [23] developed this approach further by proposing different roles alongside an algorithm for document collection that allows "one searcher to explore the information landscape broadly, looking for serendipitous discovery, while the other pursues specific highly relevant information from a particular aspect" [23](pg. 3). This approach is less constricting as roles can also be traded as a search session progresses.

Roles can be useful for collaboration as they comprise shared expectations about how each person in a group might be expected to behave [13]. In this way, roles can simplify and structure collaboration during information seeking. Yet one downside of role-based mediation is that even minor differences in the functional or descriptive assignment of roles may unintentionally imply status differences [13]. Moreover, asymmetric role assignment could be constraining in certain situations where all participants wish to participate equally in the search process. Although studies suggest that a clear set of well-defined roles can be helpful for groups, leading to positive effects on dynamics and performance [13], setting up and defining roles is an additional cost, and can lead to conflict through role ambiguity, strain, or overlapping duties between roles [13]. This suggests that lightweight or informal role assignment could be preferable. Gonzalez-Ibañez et al. [10] also noted that role-based DoL relies on user awareness of the responsibilities and abilities afforded to each role, meaning that users need to establish common ground over individual duties.

With regard to the CIS literature, prior studies of role-based DoL have only examined concurrent search scenarios [23, 11] and little is known about the benefit of roles for asynchronous work. Role-based approaches like those described by [23] only work if users do not change their roles and responsibilities during the process of collaboration [10]. Finally, although it has been shown that role assignment can have a positive impact on the results achieved, there has been no formal comparison of different role assignments and setups in CIS. The studies that do exist in this area tend to compare role-based collaboration to groups with no role assignment (e.g. [11]). Also, most evaluations have focused on pairs of participants, but there are often situations where CIS occurs between large groups. For example, in our own research on travel planning tasks, we are currently examining CIS among a group of 13 people planning a family holiday. Clearly the costs may outweigh the benefits when organising a role assignment in such a large group, leading us to wonder how effective DoL might be achieved.

3. SUMMARY AND FUTURE WORK

Through reviewing the literature on DoL in CIS, we have been able to highlight a variety of strengths and weaknesses associated with different approaches. Communicative DoL is high in user freedom but incurs the greatest coordination costs. UI-level mediation can alleviate some of these costs, but knowing which features to employ in a system (and even whether or not users will value them) is difficult. Algorithmic mediation has been shown to be beneficial but its utility may depend on the task at hand. Finally, role-based DoL quickly clarifies the roles of team members but may be too constraining for certain CIS tasks.

How should research on division of labour proceed given the issues we have highlighted? In the following subsections, we draw on our review alongside our own work to map out several areas in which future research effort might be focused. Our suggestions are of course very preliminary, and we hope to extend these areas through collective discussion at the workshop.

3.1 Comparison of Approaches

The vast majority of prior studies in CIS involve the use of a single division of labour policy. There has been almost no work that attempts to to perform comparative studies that vary the method of division of labour used to complete the same CIS task. By assessing the performance of searchers when using one division of labour approach over the other, researchers might be able to determine which approach is better for certain CIS activities. It may also be the case that a certain combination of approaches are more optimal for collaboration than others (although determining what constitutes 'optimal' collaboration is indeed a broader question).

3.2 Exploring Different Tasks

In a related vein, it is possible that users' preferred division of labour policy could be affected by the task at hand. This question is important because the intended use of a system should guide the DoL policies it includes, i.e., a collaborative travel planning system might allow searchers to divide results by price, location, and so on. Earlier in this paper, we noted that selective filtering might be more appropriate for precision-oriented searches, whereas more open-ended search activities may be hindered by this approach. Indeed, there has been little work examining how well division of labour policies can map to different information seeking tasks. In our own work, we have been conducting studies of collaborators working on real-world travel planning tasks. One finding is that searchers frequently organise their work by role assignment, yet also retrace the steps of their collaborators in an attempt to understand the broader information seeking process.

3.3 User Studies of Coordination

Much CIS research seems to be based on an *a priori* understanding of division of labour, in that there have been relatively few studies (outside of foundational work, e.g. [15]) that examine how people divide work and coordinate during Web-based CIS when they are *not* supported by technology. This is of concern given that analysis of the ways in which people naturally coordinate may allow us to design tools that reflect these tendencies. Although some studies do exist (for example, survey studies by Morris [15] and Capra et al. [3]), their findings are based on self-report rather than direct empirical observation. And, as we noted in our discussion of UI-level mediation, a majority of CIS systems have been evaluated using only short term studies. This presents two opportunities: fist, an opportunity to study naturalistic strategies employed by collaborative searchers; and second, an opportunity to assess division of labour policies, and CIS systems more generally, over the longer term using field studies conducted over periods of several weeks or months.

In our work, we have been undertaking a number of studies to address these issues. First, we have conducted a series of experiments where we ask pairs of searchers to agree a quantitative division of labour for an information-seeking task requiring 10 sources on a particular topic. Thus far we have examined popular

music, slime mold, design psychology and art crime (each pair works on one topic, and in one experiment we used topic as an independent variable). We frame the allocation procedure using a reductionist model of division of labour (see [12] for more detail) which requires searchers to come to an explicit agreement about their individual allocations. These experiments allow us to explore the issues that come to the fore when collaborators decide how to allocate work (e.g., skill, expertise, time, knowledge).

We have also examined the coordination strategies employed by searchers to prevent redundancy during the actual completion of the task. We have observed four such strategies:

- *Partition Document Space.* Here searchers explicitly coordinate work by searching for different types of sources. For example, we have seen searchers dividing work by independently searching for scientific articles versus blogs and websites.
- *Partition Web Space.* This strategy involves using different Web services to coordinate work. We have observed pairs using this strategy to partition, for example, Google scholar versus EBSCO, or Bing versus ISI Web of Knowledge.
- *Partition Semantic Space.* This strategy involves choosing different aspects of the topic, independently of web services or document types, and then dividing these aspects between collaborators. For example, we have seen that searchers working on a recall-oriented task on the topic of 'popular music' organise searches according to their own knowledge of musical history, taking into account their own artists or genres of preference.
- *Partition Keyword Space.* Finally, we have observed searchers explicitly sharing their search terms so as to avoid redundancy. For example, if one searcher used 'information seeking' as a search term, the other would use a different query in the hope of retrieving different results.

Future CIS systems might benefit from supporting these strategies given that they are employed by searchers during the completion of real-world CIS tasks. Moreover, these are likely not the only available strategies for coordinating work, and thus future research might continue to explore emergent coordination behaviour in the hope of identifying new opportunities for creating DoL.

3.4 Studies of Knowledge and Skill

Finally, we have seen that differences in knowledge and skill can affect the way in which searchers organise work. We have also seen that allocations can shift during completion of work, particularly when skill disparity becomes apparent. This leads us to question whether we might able to design CIS tools that allow for more flexible and adjustable divisions of labour. We note that prior approaches to division of labour are somewhat rigid in terms of the actual quantities of work assigned to each collaborator – allocations tend to be made on the basis of equity, with each collaborator given a similar amount of work. Yet tools that automatically allocate results might also allow the proportion of work to be shifted among team members to account for differences in skills and expertise.

4. CONCLUSION

Division of labour is often necessary for collaborative information seeking. In this paper, we reviewed four approaches to division of labour in CIS and suggested a variety of future directions. Our hope is that by attending the workshop, we will be able to provoke discussion about what is achievable, what is feasible, and what is appropriate when designing for division of labour in collaborative information seeking. In turn, we hope that further research will lead to improvements in the design of future CIS tools.

5. REFERENCES

[1] Buchanan, G., and Loizides, F. Investigating document triage on paper and electronic media. In *Research and Advanced Technology for Digital Libraries*. 2007, 416–427.

[2] Capra, R., Chen, A. T., Hawthorne, K., Arguello, J., Shaw, L., and Marchionini, G. Design and evaluation of a system to support collaborative search. In *Proc. ASIST* (2012).

[3] Capra, R., Marchionini, G., Velasco-Martin, J., and Muller, K. Tools-at-hand and learning in multi-session, collaborative search. In *Proc. CHI* (2010), 951–960.

[4] Eason, K. Division of labour and the design of systems for computer support for cooperative work. *Journal of Information Technology 11* (1996), 39–50.

[5] Foley, C., and Smeaton, A. Synchronous collaborative information retrieval: Techniques and evaluation. In *Proc. ECIR* (2009), 42–53.

[6] Foley, C., and Smeaton, A. Division of labour and sharing of knowledge for synchronous collaborative information retrieval. *Info. Proc. & Man. 46* (2010), 762–772.

[7] Golovchinsky, G., Adcock, J., Pickens, J., Pernilla, Q., and Back, M. Cerchiamo: a collaborative exploratory search tool. In *Proc. CSCW*, Demo Session (2008).

[8] Golovchinsky, G., Diriye, A., and Pickens, J. Designing for collaboration in information seeking. In *Proc. HCIR* (2011).

[9] Golovchinsky, G., and Pickens, J. A model of collaborative search. Available online, 2009. http://www.slideshare.net/geneg/a-model-of-collaborative-search.

[10] González-Ibáñez, R., Shah, C., and White, R. W. Pseudo Collaboration as a Method to Perform Selective Algorithmic Mediation in Collaborative IR Systems. In *Proc. ASIST* (2012).

[11] Imazu, M., Nakayama, S., and Joho, H. Effect of explicit roles on collaborative search in travel planning task. In *Proc. AIRS* (2011), 205–214.

[12] Kelly, R. M. An economic approach to studying division of labour in collaborative search tasks. In *Proc. British HCI*, BCS (Swinton, UK, 2011), 539–542.

[13] Levine, J. M., and Moreland, R. L. Small Groups. In *The Handbook of Social Psychology, Vol. 2*, D. T. Gilbert and S. T. Fiske, Eds. Mcgraw-Hill, Boston, MA, 1998, 415–469.

[14] Mark, G., Haake, J. M., and Streitz, N. A. Hypermedia structures and the division of labor in meeting room collaboration. In *Proc. CSCW '96* (1996), 170–179.

[15] Morris, M. R. A survey of collaborative web search practices. In *Proc. CHI '08* (2008), 1657–1660.

[16] Morris, M. R., and Horvitz, E. Searchtogether: an interface for collaborative web search. In *Proc. UIST '07* (2007), 3–12.

[17] Morris, M. R., Lombardo, J., and Wigdor, D. Wesearch: supporting collaborative search and sensemaking on a tabletop display. In *Proc. CSCW* (2010), 401–410.

[18] Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., and Back, M. Algorithmic mediation for collaborative exploratory search. In *Proc. SIGIR* (2008), 315–322.

[19] Poltrock, S., Grudin, J., Dumais, S., Fidel, R., Bruce, H., and Pejtersen, A. M. Information seeking and sharing in design teams. In *Proc. GROUP* (2003), 239–247.

[20] Shah, C., and González-Ibáñez, R. Exploring information seeking processes in collaborative search tasks. In *Proc. ASIST* (2010), 60:1–60:10.

[21] Shah, C., and González-Ibáñez, R. Evaluating the synergic effect of collaboration in information seeking. In *Proc. SIGIR* (2011), 913–922.

[22] Shah, C., Marchionini, G., and Kelly, D. Learning design principles for a collaborative information seeking system. In *Proc. CHI '09* (2009), 3419–3424.

[23] Shah, C., Pickens, J., and Golovchinsky, G. Role-based results redistribution for collaborative information retrieval. *Info. Proc. & Man. 46*, 6 (2010), 773–781.